

CLAIMS

1 1. A system for improved load balancing in a client/server environment,
2 comprising:
3 at least one caching/ hashing switch (CHS) coupled between clients and servers
4 in said client/server environment, said CHS storing previously-requested objects;
5 wherein object requests for objects stored in said CHS are satisfied
6 immediately from said CHS.

1 2. The systems of claim 1, wherein said CHS also hashes object requests, and
2 wherein:
3 object requests which are not stored in said CHS are hashed;
4 each of said hashed object requests are forwarded to a respective server on
5 which each requested object is stored;
6 each of said requested objects is forwarded to said CHS and stored thereon;
7 and
8 a copy of each of said requested objects is returned to a respective client
9 requesting said object.

1 3. The system of claim 2, wherein said objects are web objects and wherein
2 said CHS comprises:
3 a web proxy cache; and

4 a URL-hashing switch coupled to said web proxy cache.

1 4. The system of claim 2, wherein said objects are web objects and wherein
2 said CHS comprises:

3 software means configured to operate as a web proxy cache for storing
4 retrieved web objects; and

5 software means configured to operate as a URL-hashing switch, for hashing
6 said web object requests and forwarding said hashed web object requests to said
7 respective servers.

1 5. The system of claim 4, wherein said client/server environment comprises a
2 plurality of clients coupled to at least one server farm via a network connection.

3 6. The system of claim 4, wherein said client server environment comprises a
4 plurality of clients coupled to a plurality of server farms via a network connection,
5 and wherein each of said server farms has a CHS associated therewith, and wherein
6 said system further comprises:

7 a Level 4 switch coupled between said network connection and said CHS's.

1 7. A method of improved load balancing in a client/server environment,
2 comprising the steps of:

3 receiving an object request from a client;

determining if the object requested by said object request is stored in a cache coupled between said client and a server farm;

if said object is stored in said cache, immediately returning a copy of said object to said client; and

if said object is not stored in said cache, then:

hashing said object request;

forwarding said hashed object request to said server farm;

forwarding said requested object from said server farm to said cache for storage; and

returning a copy of said requested object to said client.

8. A computer program product for providing improved load balancing in a client/server environment, comprising:

means for receiving an object request from a client;

means for determining if the object requested by said object request is stored in a cache coupled between said client and a server farm;

means for immediately returning a copy of said object to said client if said object is stored in said cache; and

means for:

hashing said object request;

forwarding said hashed object request to said server farm;

11 forwarding said requested object from said server farm to said cache for
12 storage; and
13 returning a copy of said requested object to said client,
14 if said object is not stored in said cache.

1 9. An improvement to a load balancing system in a client/server environment
2 having at least one client and a plurality of servers coupled via a network connection,
3 and a hashing switch coupled between said network connection and said plurality of
4 servers, said improvement comprising:
5 a cache coupled between said network connection and said hashing switch,
6 said cache storing previously requested objects and configured to satisfy requests for
7 said previously requested objects without passing said requests to said hashing switch.